

The McLennan-Zubarev steady state distribution and fluctuation theorems

Mitsusada M. Sano

*Graduate School of Human and Environmental Studies,
Kyoto University, Yoshida-Nihonmatsu-cho, Sakyo, Kyoto, 606-8501, Japan*

Abstract

The McLennan-Zubarev steady state distribution is studied in the connection with fluctuation theorems. We derive the McLennan-Zubarev steady state distribution from the nonequilibrium detailed balance relation. Then, considering the cumulant function or cumulant functional, two fluctuation theorems for entropy and for currents are proved. Using the fluctuation theorem for currents, the current is expanded in terms of thermodynamic forces. In the lowest order of the thermodynamic force, we find that the transport coefficient satisfies the Onsager's reciprocal relation. In the next order, we derived the correction term to the Green-Kubo formula.

Keywords: Nonequilibrium steady state, McLennan-Zubarev steady state distribution, Fluctuation theorems

PACS: 05.40.-a, 05.60.-k, 05.90.+m, 05.20.-y

1. Introduction

Systems out equilibrium are filled in the nature, for instance, physical, chemical, and biological phenomena. In an isolated system, a nonequilibrium state spontaneously relaxes to thermodynamical equilibrium. On the other hand, a nonequilibrium steady state (NESS) is maintained or sustained by flows of energy and matter from the outside of a given system. Macroscopically, the fluxes and the entropy production are constant in the NESS. However, microscopically speaking, the fluxes and the entropy production fluctuate around macroscopic steady values. To elucidate the nature of these fluctuations is an issue of nonequilibrium statistical physics in last two decades, for instance, in refs. [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13].

The entropy production is the most interesting quantity in the NESS, since it measures irreversibility of a given system. The fluctuation of the entropy production has a symmetry, that is, in the NESS, the logarithm of the probability ratio of the entropy production is related to the entropy production itself.

$$\lim_{\tau \rightarrow \infty} \frac{1}{\tau} \ln \frac{P_{\tau}(\sigma)}{P_{\tau}(-\sigma)} = \sigma, \quad (1)$$

where σ is the entropy production and $P_{\tau}(\sigma)$ is the probability that the system exhibits the entropy production σ in time interval τ . This relation is called the fluctuation theorem, which was first found in numerical data of the simulations of the Nosé-Hoover thermostat systems[1]. After that, a careful examination using dynamical systems theory and ergodic theory, has elucidated the phenomena observed[2, 3]. As a result, it is shown that the fluctuation of the entropy production is governed by time reversal symmetry of the system. The probability ratio in eq. (1) is a ratio of probabilities of forward path and its time reversed path. The fluctuation theorem is the consequence of the time reversal symmetry in the steady state. As another aspect, the fluctuation theorem affects transport phenomena. In ref. [4], it was formulated as an extension of Onsager's reciprocity relation. Furthermore, the fluctuation theorem is not restricted to dynamical systems, and was also confirmed for the Langevin system[5], for general stochastic systems[6], and for the master equation[10, 13]. In ref. [6], for stochastic systems, the fluctuation theorem is rewritten in terms of the cumulant generating function $\mu(\lambda)$, i.e., $\mu(\lambda) = \mu(1 - \lambda)$. We also call this symmetry the fluctuation theorem for the entropy production or Lebowitz-Spohn symmetry. Using the nonequilibrium detailed balance relation, microscopic derivations of the fluctuation theorem were attempted for a stochastic system[7] and for a Hamiltonian

system[8]. For dynamical systems, the proof of the fluctuation theorem is still not in a satisfactory form except the case of the Gaussian thermostat systems or Nosé-Hoover systems. Here we emphasize that we treat the Hamiltonian dynamical systems, not stochastic systems. We overcome deficiency of mathematical rigor in Jarzynski's treatment[8]. As we shall see, the deficiency is due to the fact that he does not know the steady state distribution. In addition, our formulation is closely related to the work by Lebowitz and Spohn[6]. They do not derive the fluctuation theorem in the section 7, entitled "Stochastic and thermostating heat reservoirs", of [6]. Our results may be what they tried to attempt.

Independent from the fluctuation theorem, the derivation of the steady state distribution in a form like a canonical ensemble was one of themes in nonequilibrium statistical physics. It was derived by McLennan[14, 15, 16] and by Zubarev[17, 18]. Now their steady state distribution is called the McLennan-Zubarev steady state distribution. The McLennan-Zubarev steady state distribution was very formally derived. For stochastic systems (the master equation and the Langevin equation), the meaning of the McLennan-Zubarev steady state distribution is investigated by several authors[19, 20, 21]. In these works, the steady state distribution near equilibrium is written in a similar form to the McLennan-Zubarev steady state distribution. Thus, the McLennan-Zubarev steady state distribution is considered as an approximation of the steady state distribution, tentatively. However, in recent years, in the considerations of the steady state thermodynamics[22, 23, 24, 25], the McLennan-Zubarev steady state distribution was derived through another route, i.e., the nonequilibrium detailed balance relation, which is closely related to the fluctuation theorem. In addition, the McLennan-Zubarev steady state distribution was also derived for quantum systems[26, 27]. These observations strongly stimulate us to reconsider the McLennan-Zubarev steady state distribution in a connection with the fluctuation theorem.

The aim of this paper is to derive the McLennan-Zubarev steady state distribution from the nonequilibrium detailed balance relation, which is a key relation to derive the fluctuation theorem, and, in addition, to prove the fluctuation theorems for the entropy production and for the currents, by using the derived McLennan-Zubarev steady state distribution. Thus, we establish an exact connection between the McLennan-Zubarev steady state distribution and the fluctuation theorems. We confirm that the McLennan-Zubarev steady state distribution is exact, as far as we admit the nonequilibrium detailed balance relation. For stochastic Markovian dynamics, the nonequilibrium detailed balance relation is justified[28]. Unfortunately, however, we cannot justify the nonequilibrium detailed balance relation in our Hamiltonian formulation. This relation is an assumption more or less. We believe that the assumption of this relation holds. As shown later, the nonequilibrium detailed balance relation is strongly related not only to the McLennan-Zubarev steady state distribution, but also to the fluctuation theorems. The nonequilibrium detailed balance relation, eq.(53), implies all we need, i.e., McLennan-Zubarev steady state distribution and two fluctuation theorems. As done for the master equation[11, 12], we obtain correction terms to the Green-Kubo formula, that is, the non-linear response, by using the cumulant generating functional for the currents. It is sure that the derived corrections are useful for investigations on the transport property.

This paper is organized as follows. In § 2, some known results on the McLennan-Zubarev steady state distribution are summarized. In § 3, after a preparation of the setting and notations, the McLennan-Zubarev steady state distribution is derived from the nonequilibrium detailed balance relation. In § 4, using a similar discussion to that in the previous section, the fluctuation theorem for the entropy production is proved. In § 5, the fluctuation theorem for the currents is proved. In § 6, the nonequilibrium detailed balance relation is argued. In § 7, we investigate the mean current. The mean current is expressed in a power series of thermodynamic forces. We derive the linear response (i.e., the Green-Kubo formula) and the non-linear response in the lowest order. As a by-product, we obtain non-trivial relations of the current correlation functions. These non-trivial relations are a kind of representations of time-reversal symmetry. In § 8, we summarize the results of this paper.

2. The McLennan-Zubarev steady state distribution

In this section, we summarize a setting up for defining the McLennan-Zubarev steady state distribution. The basic facts on the McLennan-Zubarev steady state distribution are in [18] in details. The derivation of the McLennan-Zubarev steady state distribution, which is not the derivation in [18], will be shown in the next section.

First, we summarize known results of nonequilibrium thermodynamics. We consider an isotropic one-component

fluid. The entropy density s satisfies the following entropy balance equation [29, 30].

$$\frac{\partial s}{\partial t} + \text{div}(\mathbf{j}_s) = \sigma[s], \quad (2)$$

where \mathbf{j}_s is the entropy current

$$\mathbf{j}_s = s\mathbf{u} + \frac{\mathbf{Q}}{T} - \frac{\mu^*\mathbf{j}}{T}. \quad (3)$$

Here T is the temperature. \mathbf{u} is the local fluid velocity. \mathbf{Q} is the heat current. \mathbf{j} is the diffusion current. The first term of the right hand side of eq.(3) is the entropy current of the fluid. The second term is the entropy current of the energy flow. The third term is the entropy current of the diffusion. μ is the chemical potential. We set

$$\mu^* = \mu - \frac{\mathbf{u}^2}{2}. \quad (4)$$

The energy current \mathbf{j}_ε is given by

$$\mathbf{j}_\varepsilon = \mathbf{Q} - \mathbf{u} : \sigma'. \quad (5)$$

σ' is the viscous stress tensor. The stress tensor \mathbf{T} is related to the viscous stress tensor as follows.

$$\mathbf{T}_{ij} = -P\delta_{ij} + \sigma'_{ij}. \quad (6)$$

P is the hydrostatic pressure. The product of vector and tensor $\mathbf{u} : \sigma'$ is

$$(\mathbf{u} : \sigma')_i = \sum_j u_j \sigma'_{ji}. \quad (7)$$

The entropy production rate is given by

$$\sigma[s] = \mathbf{j}_\varepsilon \cdot \nabla \left(\frac{1}{T} \right) + (-\sigma') : \nabla \left(-\frac{\mathbf{u}}{T} \right) + \mathbf{j} \cdot \nabla \left(-\frac{\mu^*}{T} \right). \quad (8)$$

The notation of the tensor product is

$$(-\sigma') : \nabla \left(-\frac{\mathbf{u}}{T} \right) = \sum_{i,j} (-\sigma'_{ij}) \frac{\partial}{\partial x_j} \left(-\frac{u_i}{T} \right). \quad (9)$$

From the expression of $\sigma[s]$, we see that the entropy production rate is (the irreversible part of the current of the conserved quantity) \times (the gradient of the intensive parameter). Each term is related to the energy conservation law, the momentum conservation law, and the particle number conservation law. This fact is important when we consider the entropy production rate of a given system not only in a macroscopic description, but also microscopic description.

Now we turn to microscopic description of a one-component fluid system. Consider a system with N point particles with mass m . The interaction potential is given by $V(|\mathbf{q}_i - \mathbf{q}_j|)$. The Hamiltonian is

$$H = \sum_{n=1}^N \frac{\mathbf{p}_n^2}{2m} + \sum_{i<j}^N V(|\mathbf{q}_i - \mathbf{q}_j|). \quad (10)$$

The equations of motion is

$$\frac{d\mathbf{q}_i}{dt} = \frac{\partial H}{\partial \mathbf{p}_i}, \quad \frac{d\mathbf{p}_i}{dt} = -\frac{\partial H}{\partial \mathbf{q}_i}. \quad (11)$$

In this paper, we use the following notation for the variables of position and momentum. $\Gamma = (\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_N, \mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N)$. The Hamiltonian H is a function of $\{\mathbf{q}_i\}_{i=1}^N$ and $\{\mathbf{p}_i\}_{i=1}^N$. The Hamiltonian may written as $H = H(\{\mathbf{q}_i\}, \{\mathbf{p}_i\})$. Symmetry which the system possesses is important. This Hamiltonian has a symmetry.

$$H(\{\mathbf{q}_i\}, \{\mathbf{p}_i\}) = H(\{\mathbf{q}_i\}, -\{\mathbf{p}_i\}). \quad (12)$$

Time-reversal symmetry of orbits can be written as

$$\mathbf{q}_{i,R}(t) = \mathbf{q}_i(-t), \quad \mathbf{p}_{i,R}(t) = -\mathbf{p}_i(-t). \quad (13)$$

For the orbit $\Gamma(t)$, the time-reversed orbit is written as $\Gamma_R(t)$.

This system has three conservation laws. (1) Particle number conservation law:

$$\frac{\partial n(\Gamma, \mathbf{x})}{\partial t} + \text{div}(\mathbf{j}(\Gamma, \mathbf{x})) = 0. \quad (14)$$

$n(\Gamma, \mathbf{x})$ is the particle density. $\mathbf{j}(\Gamma, \mathbf{x})$ is the current density of particles. (2) Energy conservation law:

$$\frac{\partial H(\Gamma, \mathbf{x})}{\partial t} + \text{div}(\mathbf{J}_H(\Gamma, \mathbf{x})) = 0. \quad (15)$$

$H(\Gamma, \mathbf{x})$ is the energy density. $\mathbf{J}_H(\Gamma, \mathbf{x})$ is the energy current density. (3) Momentum conservation law:

$$\frac{\partial \mathbf{p}(\Gamma, \mathbf{x})}{\partial t} + \text{div}(\mathbf{T}(\Gamma, \mathbf{x})) = 0. \quad (16)$$

$\mathbf{T}(\Gamma, \mathbf{x})$ is the momentum tensor. Here the functions appeared in the above conservation laws are the density of particles,

$$n(\Gamma, \mathbf{x}) = \sum_{i=1}^N \delta(\mathbf{q}_i - \mathbf{x}), \quad (17)$$

the momentum density,

$$\mathbf{p}(\Gamma, \mathbf{x}) = \sum_{i=1}^N \mathbf{p}_i \delta(\mathbf{q}_i - \mathbf{x}), \quad (18)$$

and the current of particles,

$$\mathbf{j}(\Gamma, \mathbf{x}; t) = \sum_{i=1}^N \frac{\mathbf{p}_i}{m} \delta(\mathbf{q}_i - \mathbf{x}). \quad (19)$$

In order to capture heat, we define the energy density is given by

$$H(\Gamma, \mathbf{x}) = \sum_{i=1}^N \left\{ \frac{\mathbf{p}_i^2}{2m} + \frac{1}{2} \sum_{j \neq i} V(|\mathbf{q}_i - \mathbf{q}_j|) \right\} \delta(\mathbf{q}_i - \mathbf{x}). \quad (20)$$

For heat current, the energy current is defined as

$$\mathbf{J}_H(\Gamma, \mathbf{x}; t) = \sum_i^N \left[\frac{\mathbf{p}_i^2}{2m} + \frac{1}{2} \sum_{j \neq i} V(|\mathbf{q}_i - \mathbf{q}_j|) \right] \frac{\mathbf{p}_i}{m} \delta(\mathbf{q}_i - \mathbf{x}) + \frac{1}{4m} \sum_{j \neq i} ((\mathbf{p}_i + \mathbf{p}_j) \cdot \mathbf{F}_{ij})(\mathbf{q}_i - \mathbf{q}_j) \delta(\mathbf{q}_i - \mathbf{x}). \quad (21)$$

The momentum tensor $\mathbf{T}(\Gamma, \mathbf{x}; t)$ is

$$T_{\beta\alpha}(\Gamma, \mathbf{x}; t) = \frac{1}{2} \sum_{i \neq j} (\mathbf{q}_i - \mathbf{q}_j)_\beta (F_{ij})_\alpha \delta(\mathbf{q}_i - \mathbf{x}) + \frac{1}{m} \sum_{i=1}^N (\mathbf{p}_i)_\beta (\mathbf{p}_i)_\alpha \delta(\mathbf{q}_i - \mathbf{x}), \quad (22)$$

Here the force is

$$\mathbf{F}_{ij} = -\frac{\partial}{\partial \mathbf{q}_i} V(|\mathbf{q}_i - \mathbf{q}_j|). \quad (23)$$

$\mathbf{F}_{ij}(\mathbf{q}_i - \mathbf{q}_j)$ is the force, which the j -th particle acts to the i -th particle

As a result, the above three conservation law implies three phenomena, i.e., thermal conduction, momentum diffusion, and diffusion. Here we define three currents corresponding to the thermal conduction, momentum diffusion, and diffusion.

$$\mathbf{j}_0(\Gamma, \mathbf{x}; t) = \mathbf{J}_H(\Gamma, \mathbf{x}; t) \quad (24)$$

$$\mathbf{j}_1(\Gamma, \mathbf{x}; t) = \mathbf{T}(\Gamma, \mathbf{x}; t) \quad (25)$$

$$\mathbf{j}_2(\Gamma, \mathbf{x}; t) = \mathbf{j}(\Gamma, \mathbf{x}; t). \quad (26)$$

Next we define three quantities.

$$P_0(\Gamma, \mathbf{x}) = H(\Gamma, \mathbf{x}) \quad (27)$$

$$P_1(\Gamma, \mathbf{x}) = \mathbf{p}(\Gamma, \mathbf{x}) \quad (28)$$

$$P_2(\Gamma, \mathbf{x}) = n(\Gamma, \mathbf{x}). \quad (29)$$

Here μ^* is the chemical potential, which is defined as

$$\mu^* = \mu - \frac{m\mathbf{u}^2}{2}, \quad (30)$$

where \mathbf{u} is the velocity field. If the velocity field exists, the chemical potential would be subtracted by the amount of the velocity field. $\beta(\mathbf{x})$ is the inverse temperature divided by the Boltzmann constant.

$$\beta(\mathbf{x}) = \frac{1}{k_B T(\mathbf{x})}, \quad (31)$$

Here $T(\mathbf{x})$ is the temperature field. We also define the following quantities.

$$F_0(\mathbf{x}) = \beta(\mathbf{x}) \quad (32)$$

$$F_1(\mathbf{x}) = -\beta(\mathbf{x})\mathbf{u}(\mathbf{x}) \quad (33)$$

$$F_2(\mathbf{x}) = -\beta(\mathbf{x})\mu^*(\mathbf{x}). \quad (34)$$

The thermodynamic forces $\mathbf{X}_i = \nabla F_i$ are defined as

$$\mathbf{X}_0(\mathbf{x}) = \nabla\beta(\mathbf{x}) \quad (35)$$

$$\mathbf{X}_1(\mathbf{x}) = -\nabla(\beta(\mathbf{x})\mathbf{u}) \quad (36)$$

$$\mathbf{X}_2(\mathbf{x}) = -\nabla(\beta(\mathbf{x})\mu^*). \quad (37)$$

For convenience, we define a vector, which combines three thermodynamic forces.

$$\mathbf{X}(\mathbf{x}) = \begin{pmatrix} \mathbf{X}_0(\mathbf{x}) \\ \mathbf{X}_1(\mathbf{x}) \\ \mathbf{X}_2(\mathbf{x}) \end{pmatrix}. \quad (38)$$

In nonequilibrium thermodynamics, the local entropy production rate $\sigma[s]$ is defined as eq.(8). Now we define the microscopic local entropy production rate $\sigma[\Gamma, \mathbf{x}]$ for our particle system.

$$\begin{aligned} \sigma[\Gamma, \mathbf{x}] &= \sum_m \mathbf{j}_m(\Gamma, \mathbf{x}; t) \cdot \mathbf{X}_m(\mathbf{x}) \\ &= \mathbf{J}_H(\Gamma, \mathbf{x}; t) \cdot \nabla\beta(\mathbf{x}) + \mathbf{T}(\Gamma, \mathbf{x}; t) : \nabla(-\beta(\mathbf{x})\mathbf{u}) + \mathbf{j}(\Gamma, \mathbf{x}; t) \cdot \nabla(-\beta(\mathbf{x})\mu^*). \end{aligned} \quad (39)$$

Precisely speaking, this quantity is the local entropy production rate divided by the Boltzmann constant k_B . Equation (39) is a first key assumption in this paper. Other assumptions are the nonequilibrium detailed balance relation, and the time-reversal symmetry, which will be explained later. These three assumptions are the key relations in this paper.

The local equilibrium state is characterized by a local equilibrium distribution. That is

$$\rho_\ell(\Gamma) = Z_0^{-1} \exp \left[- \sum_m \int d\mathbf{x} F_m(\mathbf{x}) P_m(\Gamma, \mathbf{x}) \right], \quad (40)$$

where

$$Z_0 = \int d\Gamma \exp \left[- \sum_m \int d\mathbf{x} F_m(\mathbf{x}) P_m(\Gamma, \mathbf{x}) \right]. \quad (41)$$

ℓ stands for local equilibrium. This distribution does not describe a steady state. To obtain the steady state distribution, we should treat the deviation from the local equilibrium. The deviation includes the time-integral of the entropy production rate.

Here we “simply” write down the McLennan-Zubarev steady state distribution which appeared in the book of Zubarev[18]. By the above setting, the McLennan-Zubarev steady state distribution is given as follows[14, 15, 16, 17, 18].

$$\rho_{ss}(\Gamma) = Z^{-1} \exp \left[- \sum_m \int d\mathbf{x} F_m(\mathbf{x}) P_m(\Gamma, \mathbf{x}) + \sum_m \int d\mathbf{x} \int_{-\infty}^0 dt \mathbf{j}_m(\Gamma, \mathbf{x}; t) \cdot \mathbf{X}_m(\mathbf{x}) \right]. \quad (42)$$

“ss” means the steady state. Here Γ is $\Gamma = (\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_N, \mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N)$. Z is the normalization constant.

$$Z = \int d\Gamma \exp \left[- \sum_m \int d\mathbf{x} F_m(\mathbf{x}) P_m(\Gamma, \mathbf{x}) + \sum_m \int d\mathbf{x} \int_{-\infty}^0 dt \mathbf{j}_m(\Gamma, \mathbf{x}; t) \cdot \mathbf{X}_m(\mathbf{x}) \right]. \quad (43)$$

$\mathbf{j}_m(\Gamma, \mathbf{x}; t)$ is the energy current density ($m = 0$), the momentum current density ($m = 1$), and the current density of the particle number ($m = 2$). It might be that an explanation on the notations is needed here. The coordinates of Γ and \mathbf{x} is the coordinate of the orbits and of the density, respectively. $\mathbf{j}_m(\Gamma, \mathbf{x}; t)$ is the density of the current. The McLennan-Zubarev steady state distribution is near the local equilibrium state, eq.(40). The deviation from the local equilibrium is

$$\sum_m \int d\mathbf{x} \int_{-\infty}^0 dt \mathbf{j}_m(\Gamma, \mathbf{x}; t) \cdot \mathbf{X}_m(\mathbf{x}). \quad (44)$$

This term of eq. (42) expresses the total amount of the entropy production for a given orbit. $\rho_{ss}(\Gamma)$ has a physical meaning that at infinite past ($t = -\infty$), the system is in a local equilibrium. Then after, up to the present ($t = 0$), the system produces the whole of the entropy production. This amount of the total entropy production is in the exponential function.

3. Derivation of the McLennan-Zubarev steady state distribution

In this section, we derive the McLennan-Zubarev steady state distribution from the nonequilibrium detailed balance relation. Assumptions are needed to derive the McLennan-Zubarev steady state distribution: (a) the entropy production rate, eq.(39), (b) the time-reversibility, eq.(57), and (c) the nonequilibrium detailed balance relation, eq.(59). The nonequilibrium detailed balance relation is needed to derive the McLennan-Zubarev steady state distribution and the fluctuation theorems. The time reversibility is essential to the fluctuation theorems. Before starting the derivation, we prepare a setting and notations for later use.

3.1. Preparation

In this subsection, we present the notations and the time-reversal symmetry, which are important for our formulation. Here we do not follow McLennan’s derivation, but will present another derivation by using the nonequilibrium detailed balance relation. The McLennan’s derivation is presented in Appendix A for comparison.

Consider a system which consists of particles. The total system has the system and two baths. The Hamiltonian of the total system is

$$H_{tot} = H_S + H_B + U. \quad (45)$$

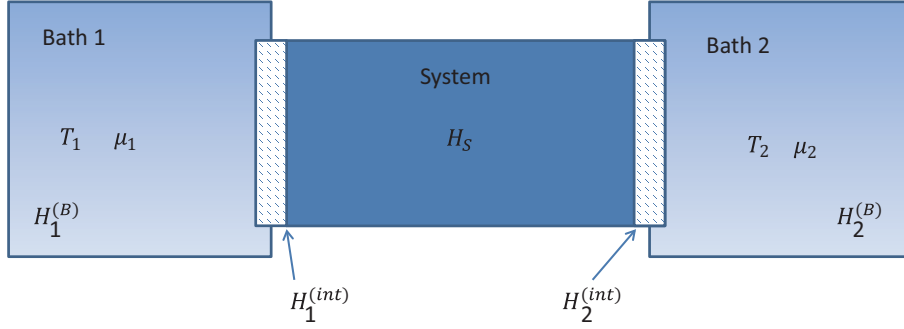


Figure 1: The configuration of the system and the baths: The system is sandwiched between the baths 1 and 2, whose temperatures and chemical potentials are different ($T_1 \neq T_2$, $\mu_1 \neq \mu_2$). Hamiltonians H_S , $H_i^{(B)}$, and $H_i^{(int)}$ ($i = 1, 2$) describe the system, the i -th bath, and the interaction between the system and the i -th bath, respectively

H_S is the Hamiltonian of the system. H_B is the Hamiltonian of the baths. U is the interaction between the system and the baths. The configuration of the whole system is depicted in Fig.1. The bath i is in equilibrium and has temperature T_i ($i = 1, 2$), and has chemical potential μ_i ($i = 1, 2$). The system (the bath i) has N_s (N_i) particles, respectively, $N = N_s + N_1 + N_2$. The total system has N particles. All N particles are identical. Imagine that the system is under the heat conduction and the particle flow with this nonequilibrium boundary condition. We define the Hamiltonian of the total system as follows.

$$H_S = H_S(\Gamma_s), \quad (46)$$

$$H_B = \sum_{i=1}^2 H_i^{(B)}(\Gamma_b^{(i)}), \quad (47)$$

$$U = \sum_{i=1}^2 H_i^{(int)}(\Gamma_b^{(i)}, \Gamma_s). \quad (48)$$

Γ_s is the coordinates and momenta of the system. $\Gamma_b^{(i)}$ is the coordinates and momenta of the bath i . $H_i^{(B)}$ is the Hamiltonian of the bath i . $H_i^{(int)}$ represents the interaction between the system and the bath i . The state of the system and the baths is determined by the coordinates and the momenta.

$$\Gamma_s = (\mathbf{q}_{1,s}, \dots, \mathbf{q}_{N_s,s}, \mathbf{p}_{1,s}, \dots, \mathbf{p}_{N_s,s}), \quad (49)$$

$$\Gamma_b^{(i)} = (\mathbf{q}_{1,i}, \dots, \mathbf{q}_{N_i,i}, \mathbf{p}_{1,i}, \dots, \mathbf{p}_{N_i,i}). \quad (50)$$

The Hamiltonians are defined as

$$H_S(\Gamma_s) = \sum_{j=1}^{N_s} \frac{|\mathbf{p}_{j,s}|^2}{2m} + V_s(\mathbf{q}_{1,s}, \dots, \mathbf{q}_{N_s,s}), \quad (51)$$

$$H_i^{(B)}(\Gamma_b^{(i)}) = \sum_{j=1}^{N_i} \frac{|\mathbf{p}_{j,b}|^2}{2m} + V_i(\mathbf{q}_{1,i}, \dots, \mathbf{q}_{N_i,i}), \quad (52)$$

$$H_i^{(int)}(\Gamma_b^{(i)}, \Gamma_s) = U_{i,s}(\mathbf{q}_{1,i}, \dots, \mathbf{q}_{N_i,i}, \mathbf{q}_{1,s}, \dots, \mathbf{q}_{N_s,s}). \quad (53)$$

We write $d\Gamma_s = \prod_{i=1}^{N_s} d\mathbf{q}_{i,s} d\mathbf{p}_{i,s}$ and $d\Gamma_b^{(i)} = \prod_{j=1}^{N_i} d\mathbf{q}_{j,i} d\mathbf{p}_{j,i}$. It is clear that the system is driven out of equilibrium with the nonequilibrium boundary conditions $T_1 \neq T_2$ and $\mu_1 \neq \mu_2$. In the system, the heat conduction and the particle flow are achieved.

The time-reversal operation is essential to the consideration of the fluctuation theorem. Flipping the sign of the momenta is a basic operation. We define the $*$ -operation by

$$\Gamma_s^* = (\mathbf{q}_{1,s}, \mathbf{q}_{2,s}, \dots, \mathbf{q}_{N_s,s}, -\mathbf{p}_{1,s}, -\mathbf{p}_{2,s}, \dots, -\mathbf{p}_{N_s,s}). \quad (54)$$

The “whole” system considered in this paper is a Hamiltonian dynamical system. But the system is a certain dynamical system with dissipation by the existence of the baths or reservoirs. The Hamiltonian does not contain the magnetic field and has the symmetry of eq. (12). We denote the whole phase space by \mathcal{S} . \mathcal{S} is $\mathcal{S} = \mathcal{S}_c^s \times \mathcal{S}_m^s \times \mathcal{S}_c^{b(1)} \times \mathcal{S}_m^{b(1)} \times \mathcal{S}_c^{b(2)} \times \mathcal{S}_m^{b(2)} = \mathcal{S}_s \times \mathcal{S}_b$, where $\mathcal{S}_c^s (\mathcal{S}_c^{b(i)})$ is the coordinate space of the system (the bath i), respectively. and $\mathcal{S}_m^s (\mathcal{S}_m^{b(i)})$ is the momentum space of the system (the bath i), respectively. We denote the flow of this whole system by Φ . If we use the notation of $\Gamma(t) = (\Gamma_s, \Gamma_b^{(1)}, \Gamma_b^{(2)})$ for the state, the flow acts as

$$\Gamma(t + \tau) = \Phi^\tau \Gamma(t). \quad (55)$$

We can write the time-evolution of the distribution using the Perron-Frobenius operator.

$$\rho(\gamma; t + \tau) = \int d\Gamma \delta(\Phi^\tau \Gamma - \gamma) \rho(\Gamma; t). \quad (56)$$

We call the delta function in the above equation the integral kernel of the Perron-Frobenius operator. This type of delta functions will be frequently used later. In this paper, we may fix τ in Φ^τ . Then, after some manipulations, we shall take the limit $\tau \rightarrow \infty$. Thus, it is convenient to define the map \mathcal{T} as $\mathcal{T} = \Phi^\tau$. By the map which the whole map \mathcal{T} is restricted to the system, the phase space volume element of the system contracts.

The map \mathcal{T} satisfies the following time reversal symmetry.

$$(\mathcal{T}(\Gamma))^* = \mathcal{T}^{-1}(\Gamma^*), \quad (57)$$

for any $\Gamma \in \mathcal{S}$. This symmetry will be used later and is the second key relation to derive the McLennan-Zubarev steady state distribution and to prove the fluctuation theorems.

3.2. Derivation

We denote the entropy production contribution from a specific path $\nu \rightarrow \gamma$, by $\Sigma[\nu \rightarrow \gamma]$, where $\nu, \gamma \in \mathcal{S}_s$. This quantity is given by

$$\Sigma[\nu \rightarrow \gamma] = \sum_m \int d\mathbf{x} \int_{-\tau}^0 dt \mathbf{j}_m(\Gamma_s, \mathbf{x}; t) \cdot \mathbf{X}_m(\mathbf{x}), \quad (58)$$

where $\gamma = (\mathcal{T}(\Gamma))_s$ with $\Gamma_s = \nu$. The integral of \mathbf{x} is taken over the system. Hereafter without specifying, $\int d\mathbf{x}$ represents the integral over the system, not over the whole system. The third key relation in our derivation is the nonequilibrium detailed balance relation. For various derivations of the fluctuation theorem, the nonequilibrium detailed balance relation was used[7]. The nonequilibrium detailed balance relation for our case is given by

$$P[\nu \rightarrow \gamma] = e^{\Sigma[\nu \rightarrow \gamma]} P[\gamma^* \rightarrow \nu^*], \quad (59)$$

where $P[\nu \rightarrow \gamma]$ is the kernel of the Perron-Frobenius operator which casts from the state ν to the state γ . Since we consider a dynamical system, the function $P[\nu \rightarrow \gamma]$ is given by a delta function.

$$P[\nu \rightarrow \gamma] = \delta((\mathcal{T}(\Gamma))_s - \gamma). \quad (60)$$

Equation (59) is our assumption, since we cannot derive eq. (59) from the equations of motion by first principle. However, the nonequilibrium detailed balance relation gives us the connection between entropy production rate and the stability of the system [2, 3]. This can be easily confirmed by rewriting δ -function and integrating eq. (59) with respect to $\nu = \Gamma_s$. We get

$$\left| \frac{\partial(\mathcal{T}(\Gamma))_s}{\partial \Gamma_s} \right|^{-1} = e^{\Sigma[\nu \rightarrow \gamma]}. \quad (61)$$

Thus, the entropy production is directly related to the stability of a given system. Here we set $\nu = \Gamma_s$, $\Gamma' = \mathcal{T}(\Gamma)$, $\gamma = (\Gamma')_s$, and $\Gamma'' = \mathcal{T}^{-1}(\Gamma)$.

Let us consider time evolution of a given state. We assume that the initial state is ν and the initial distribution is in a canonical state, i.e., a local equilibrium state. This can be realized by defining the distribution $P_\nu(\Gamma_s)$

$$P_\nu(\Gamma_s) = \rho_\ell(\Gamma_s), \quad (62)$$

where

$$\rho_\ell(\Gamma) = \exp \left[- \int d\mathbf{x} \{ \beta(\mathbf{x}) H(\Gamma, \mathbf{x}) - \beta(\mathbf{x}) \mathbf{u}(\mathbf{x}) \cdot \mathbf{p}(\Gamma, \mathbf{x}) - \beta(\mathbf{x}) n(\Gamma, \mathbf{x}) \mu^*(\mathbf{x}) \} \right]. \quad (63)$$

In $P_\nu(\Gamma_s)$, ν represents the initial state. After time evolution during time interval τ , the system will be found in the state $\gamma \in \mathcal{S}_s$ with the probability

$$\rho_\nu(\gamma) = \int d\Gamma_s P_\nu(\Gamma_s) \delta((\mathcal{T}(\Gamma))_s - \gamma). \quad (64)$$

Let us start the derivation of the McLennan-Zubarev steady state distribution. Using the nonequilibrium detailed balance relation, eq. (59), we get

$$\begin{aligned} \rho_\nu(\gamma) &= \int d\Gamma_s P_\nu(\Gamma_s) P[\nu \rightarrow \gamma] \\ &= \int d\Gamma_s P_\nu(\Gamma_s) e^{\Sigma[\nu \rightarrow \gamma]} P[\gamma^* \rightarrow \nu^*] \\ &= \int d\Gamma_s P_\nu(\Gamma_s) e^{\Sigma[\nu \rightarrow \gamma]} \delta((\mathcal{T}(\mathcal{T}(\Gamma)^*))_s - \nu^*). \end{aligned} \quad (65)$$

Here we note that $\mathcal{T}(\Gamma)^* = (\mathcal{T}^{-1}(\Gamma^*))$. Thus, we obtain

$$\begin{aligned} \rho_\nu(\gamma) &= \int d\Gamma_s \rho_\ell(\Gamma_s) e^{\Sigma[\nu \rightarrow \gamma]} \delta(\Gamma_s^* - \nu^*) \\ &= \int d\Gamma_s \rho_\ell(\Gamma_s) e^{\Sigma[\nu \rightarrow \gamma]} \delta(\Gamma_s - \nu) \\ &= \rho_\ell(\nu) e^{\Sigma[\nu \rightarrow \gamma]}. \end{aligned} \quad (66)$$

We note that taking the limit $\tau \rightarrow \infty$, the steady state is obtained as

$$\rho_{ss}(\gamma) = \lim_{\tau \rightarrow \infty} Z_\tau^{-1} \rho_\nu(\gamma). \quad (67)$$

where Z_τ is the normalization constant, which is given by

$$Z_\tau = \int d\Gamma_s \exp \left[- \sum_m \int d\mathbf{x} F_m(\mathbf{x}) P_m(\Gamma_s, \mathbf{x}) + \sum_m \int d\mathbf{x} \int_{-\tau}^0 dt \mathbf{j}_m(\Gamma_s, \mathbf{x}; t) \cdot \mathbf{X}_m(\mathbf{x}) \right]. \quad (68)$$

Inserting the expression for $\rho_\nu(\gamma)$ into Eq.(67), we find

$$\rho_{ss}(\gamma) = \lim_{\tau \rightarrow \infty} Z_\tau^{-1} \rho_\ell(\nu) e^{\Sigma[\nu \rightarrow \gamma]}. \quad (69)$$

Thus, from the nonequilibrium detailed balance relation, eq. (59), we have derived the McLennan-Zubarev steady state distribution, eq. (42).

4. Fluctuation theorem for the entropy production

In this section, we prove the fluctuation theorem for the entropy production as a symmetry relation of a cumulant generating function. Now define the cumulant generating function $\mu(\lambda)$.

$$\mu(\lambda) = \lim_{\tau \rightarrow \infty} -\frac{1}{\tau} \ln \left\langle \exp \left[-\lambda \sum_m \int d\mathbf{x} \int_0^\tau dt \mathbf{j}_m(\Gamma_s, \mathbf{x}; t) \cdot \mathbf{X}_m(\mathbf{x}) \right] \right\rangle_{ss}. \quad (70)$$

Here $\langle \cdots \rangle_{ss}$ is the average over the McLennan-Zubarev steady state distribution. Note that the McLennan-Zubarev steady state distribution includes the normalization constant Z . We shall prove that this cumulant generating function has the following symmetry, which Lebowitz and Spohn discovered for stochastic systems[6],

$$\mu(\lambda) = \mu(1 - \lambda). \quad (71)$$

We call this relation the Lebowitz-Spohn symmetry.

First, we define the probability that the entropy production is X along a path $\nu \rightarrow \gamma$ during the time interval τ by $p_\tau^F(X; \gamma)$.

$$p_\tau^F(X; \gamma) = \int d\Gamma_s P_\nu(\Gamma_s) \delta((\mathcal{T}(\Gamma))_s - \gamma) \delta(X - \Sigma[\Gamma_s \rightarrow (\mathcal{T}(\Gamma))_s]). \quad (72)$$

Next consider the probability that the entropy production is X during the time interval τ . We denote this probability by $p_\tau^F(X)$. $p_\tau^F(X)$ is given by integrating eq. (72) with respect to γ . Thus, we have

$$\begin{aligned} p_\tau^F(X) &= \int d\gamma p_\tau^F(X; \gamma) \\ &= \int d\gamma \int d\Gamma_s \rho_\ell(\Gamma_s) \delta((\mathcal{T}(\Gamma))_s - \gamma) \delta(X - \Sigma[\Gamma_s \rightarrow (\mathcal{T}(\Gamma))_s]) \\ &= \int d\Gamma_s \rho_\ell(\Gamma_s) \delta(X - \Sigma[\Gamma_s \rightarrow (\mathcal{T}(\Gamma))_s]) \\ &= \int d\Gamma_s \rho_\ell(\Gamma_s) e^{\Sigma[\Gamma_s \rightarrow (\mathcal{T}(\Gamma))_s]} e^{-\Sigma[\Gamma_s \rightarrow (\mathcal{T}(\Gamma))_s]} \delta(X - \Sigma[\Gamma_s \rightarrow (\mathcal{T}(\Gamma))_s]). \end{aligned} \quad (73)$$

Inserting $\mathcal{T}\mathcal{T}^{-1}$ in the above equation, we get

$$\begin{aligned} p_\tau^F(X) &= \int d\Gamma_s \rho_\ell(\Gamma_s) \mathcal{T}\mathcal{T}^{-1} e^{\Sigma[\Gamma_s \rightarrow (\mathcal{T}(\Gamma))_s]} e^{-\Sigma[\Gamma_s \rightarrow (\mathcal{T}(\Gamma))_s]} \delta(X - \Sigma[\Gamma_s \rightarrow (\mathcal{T}(\Gamma))_s]) \\ &= \int d\Gamma_s \rho_\ell((\mathcal{T}^{-1}(\Gamma))_s) e^{\Sigma[(\mathcal{T}^{-1}(\Gamma))_s \rightarrow \Gamma_s]} e^{-\Sigma[(\mathcal{T}^{-1}(\Gamma))_s \rightarrow \Gamma_s]} \delta(X + \Sigma[(\mathcal{T}^{-1}(\Gamma))_s \rightarrow \Gamma_s]). \end{aligned} \quad (74)$$

Here we have used $\mathcal{T}^{-1}(\Sigma[\Gamma_s \rightarrow (\mathcal{T}(\Gamma))_s]) = -\Sigma[(\mathcal{T}^{-1}(\Gamma))_s \rightarrow \Gamma_s]$. In the last line, we have used the result of the Appendix B (i.e., anti-linearity).

$$\int d\Gamma F(\Gamma) \mathcal{T}G(\Gamma) = \int d\Gamma G(\Gamma) F(\mathcal{T}^{-1}(\Gamma)). \quad (75)$$

Taking $\lim_{\tau \rightarrow \infty} \int dX e^{-\lambda X}$ from the left hand side, we get

$$\begin{aligned} \lim_{\tau \rightarrow \infty} \int dX e^{-\lambda X} p_\tau^F(X) &= \lim_{\tau \rightarrow \infty} \int dX e^{-\lambda X} \int d\Gamma_s \rho_\ell((\mathcal{T}^{-1}(\Gamma))_s) \\ &\quad \times e^{\Sigma[(\mathcal{T}^{-1}(\Gamma))_s \rightarrow \Gamma_s]} e^{-\Sigma[(\mathcal{T}^{-1}(\Gamma))_s \rightarrow \Gamma_s]} \delta(X + \Sigma[(\mathcal{T}^{-1}(\Gamma))_s \rightarrow \Gamma_s]) \\ &= \lim_{\tau \rightarrow \infty} \int d\Gamma_s \rho_\ell((\mathcal{T}^{-1}(\Gamma))_s) e^{\Sigma[(\mathcal{T}^{-1}(\Gamma))_s \rightarrow \Gamma_s]} e^{-\Sigma[(\mathcal{T}^{-1}(\Gamma))_s \rightarrow \Gamma_s]} e^{\lambda \Sigma[(\mathcal{T}^{-1}(\Gamma))_s \rightarrow \Gamma_s]} \\ &= \lim_{\tau \rightarrow \infty} Z \int d\Gamma_s \rho_{ss}(\Gamma_s) e^{-(1-\lambda)\Sigma[(\mathcal{T}^{-1}(\Gamma))_s \rightarrow \Gamma_s]}. \end{aligned} \quad (76)$$

In the last line, we have used eq. (69). Thus, we obtain

$$\begin{aligned} \lim_{\tau \rightarrow \infty} \int dX e^{-\lambda X} p_\tau^F(X) &= \lim_{\tau \rightarrow \infty} Z \left\langle \exp \left[-(1-\lambda) \sum_m \int d\mathbf{x} \int_{-\tau}^0 dt \mathbf{j}_m(\Gamma_s, \mathbf{x}; t) \cdot \mathbf{X}_m(\mathbf{x}) \right] \right\rangle_{ss} \\ &= \lim_{\tau \rightarrow \infty} Z \left\langle \exp \left[-(1-\lambda) \sum_m \int d\mathbf{x} \int_0^\tau dt \mathbf{j}_m(\Gamma_s, \mathbf{x}; t) \cdot \mathbf{X}_m(\mathbf{x}) \right] \right\rangle_{ss}. \end{aligned} \quad (77)$$

In the last line, we have used the stationarity of the process, namely we can shift the integral domain in the time integral.

On the other hand, going back to eq. (73), and using the nonequilibrium detailed balance relation, eq. (59), we get

$$\begin{aligned}
p_\tau^F(X) &= \int d\gamma \int d\Gamma_s P_\gamma(\Gamma_s) \delta((\mathcal{T}(\Gamma))_s - \gamma) \delta(X - \Sigma[\Gamma_s \rightarrow (\mathcal{T}(\Gamma))_s]) \\
&= \int d\gamma \int d\Gamma_s P_\gamma(\Gamma_s) \delta(\mathcal{T}(\mathcal{T}(\Gamma)^*)_s - \gamma^*) e^{\Sigma[\Gamma_s \rightarrow (\mathcal{T}(\Gamma))_s]} \delta(X - \Sigma[\Gamma_s \rightarrow (\mathcal{T}(\Gamma))_s]) \\
&= \int d\gamma \int d\Gamma_s \rho_\ell(\Gamma_s) \delta(\mathcal{T}(\mathcal{T}(\Gamma)^*)_s - \gamma^*) e^{\Sigma[\Gamma_s \rightarrow (\mathcal{T}(\Gamma))_s]} \delta(X - \Sigma[\Gamma_s \rightarrow (\mathcal{T}(\Gamma))_s]).
\end{aligned} \tag{78}$$

Using the relation $\mathcal{T}(\mathcal{T}(\Gamma)^*) = \mathcal{T}((\mathcal{T}^{-1}(\Gamma^*))) = \Gamma^*$, we have

$$\begin{aligned}
p_\tau^F(X) &= \int d\gamma \int d\Gamma_s \rho_\ell(\Gamma_s) \delta(\Gamma_s^* - \gamma^*) e^{\Sigma[\Gamma_s \rightarrow (\mathcal{T}(\Gamma))_s]} \delta(X - \Sigma[\Gamma_s \rightarrow (\mathcal{T}(\Gamma))_s]) \\
&= \int d\gamma \int d\Gamma_s \rho_\ell(\Gamma_s) \delta(\Gamma_s - \gamma) e^{\Sigma[\Gamma_s \rightarrow (\mathcal{T}(\Gamma))_s]} \delta(X - \Sigma[\Gamma_s \rightarrow (\mathcal{T}(\Gamma))_s]) \\
&= \int d\gamma \rho_\ell(\gamma) e^{\Sigma[\gamma \rightarrow \gamma]} \delta(X - \Sigma[\gamma \rightarrow (\Gamma')_s]).
\end{aligned} \tag{79}$$

We take $\lim_{\tau \rightarrow \infty} \int dX e^{-\lambda X}$ from the left hand side and note that $\gamma = (\Gamma')_s = (\mathcal{T}(\Gamma))_s$. Then we get

$$\begin{aligned}
\lim_{\tau \rightarrow \infty} \int dX e^{-\lambda X} p_\tau^F(X) &= \lim_{\tau \rightarrow \infty} \int dX e^{-\lambda X} \int d\gamma \rho_\ell(\gamma) e^{\Sigma[\gamma \rightarrow \gamma]} \delta(X - \Sigma[\gamma \rightarrow \gamma]) \\
&= \lim_{\tau \rightarrow \infty} \int d\gamma \rho_\ell(\gamma) e^{\Sigma[\gamma \rightarrow \gamma]} e^{-\lambda \Sigma[\gamma \rightarrow \gamma]} \\
&= \lim_{\tau \rightarrow \infty} Z \int d\gamma \rho_{ss}(\gamma) e^{-\lambda \Sigma[\gamma \rightarrow \gamma]}.
\end{aligned} \tag{80}$$

In the last line, we used eq. (69). Thus, we have

$$\begin{aligned}
\lim_{\tau \rightarrow \infty} \int dX e^{-\lambda X} p_\tau^F(X) &= \lim_{\tau \rightarrow \infty} Z \left\langle \exp \left[-\lambda \sum_m \int d\mathbf{x} \int_{-\tau}^0 dt' \mathbf{j}_m(\Gamma_s, \mathbf{x}; t') \cdot \mathbf{X}_m(\mathbf{x}) \right] \right\rangle_{ss} \\
&= \lim_{\tau \rightarrow \infty} Z \left\langle \exp \left[-\lambda \sum_m \int d\mathbf{x} \int_0^\tau dt' \mathbf{j}_m(\Gamma_s, \mathbf{x}; t') \cdot \mathbf{X}_m(\mathbf{x}) \right] \right\rangle_{ss}.
\end{aligned} \tag{81}$$

In the last line, we have used the stationarity of the process. Equations (77) and (81) imply the Lebowitz-Spohn symmetry $\mu(\lambda) = \mu(1 - \lambda)$, which is the desired result.

5. Fluctuation theorem for the current

In this section, we prove the fluctuation theorem for the currents for Hamiltonian dynamical systems. For the master equation, the fluctuation theorem for the currents was proved by Andrieux and Gaspard in ref. [13].

We define the cumulant generating functional.

$$\mathcal{Q}(\lambda; \mathbf{X}) = \lim_{\tau \rightarrow \infty} -\frac{1}{\tau} \ln \left\langle \exp \left[-\sum_m \int d\mathbf{x} \int_0^\tau dt \mathbf{j}_m(\Gamma_s, \mathbf{x}; t) \cdot \lambda_m(\mathbf{x}) \right] \right\rangle_{ss}. \tag{82}$$

Note that the dependence on the thermodynamic forces $\mathbf{X}(\mathbf{x})$ comes from the McLennan-Zubarev steady state distribution. We may use the following notation.

$$\lambda(\mathbf{x}) = \begin{pmatrix} \lambda_0(\mathbf{x}) \\ \lambda_1(\mathbf{x}) \\ \lambda_2(\mathbf{x}) \end{pmatrix}. \tag{83}$$

Our target is

$$Q(\lambda; \mathbf{X}) = Q(\mathbf{X} - \lambda; \mathbf{X}), \quad (84)$$

which is the fluctuation theorem for the currents. We call this relation the Andrieux-Gaspard symmetry.

First, we define the following quantity.

$$\Xi[v \rightarrow \gamma] = \sum_m \int d\mathbf{x} \int_{-\tau}^0 dt \mathbf{j}_m(\Gamma, \mathbf{x}; t) \cdot \lambda_m(\mathbf{x}). \quad (85)$$

We are interested in how the system has the distribution of Ξ in time evolution. Thus, we define the probability that Ξ equals Y along a path $v \rightarrow \gamma$ during the time interval τ by $P_\tau^F(Y; \gamma)$. $P_\tau^F(Y; \gamma)$ is given by

$$\begin{aligned} P_\tau^F(Y; \gamma) &= \int d\Gamma_s P_v(\Gamma_s) \delta((\mathcal{T}(\Gamma))_s - \gamma) \delta(Y - \Xi[\Gamma_s \rightarrow (\mathcal{T}(\Gamma))_s]) \\ &= \int d\Gamma_s \rho_\ell(\Gamma_s) \delta((\mathcal{T}(\Gamma))_s - \gamma) \delta(Y - \Xi[\Gamma_s \rightarrow (\mathcal{T}(\Gamma))_s]). \end{aligned} \quad (86)$$

Next, we consider the probability that Ξ is Y during the time interval τ . We denote this probability by $P_\tau^F(Y)$. $P_\tau^F(Y)$ is given by integrating eq. (86) with respect to γ . Thus, we have

$$\begin{aligned} P_\tau^F(Y) &= \int d\gamma \int d\Gamma_s P_\tau^F(Y; \gamma) \\ &= \int d\gamma \int d\Gamma_s \rho_\ell(\Gamma_s) \delta((\mathcal{T}(\Gamma))_s - \gamma) \delta(Y - \Xi[\Gamma_s \rightarrow (\mathcal{T}(\Gamma))_s]) \\ &= \int d\Gamma_s \rho_\ell(\Gamma_s) \delta(Y - \Xi[\Gamma_s \rightarrow (\mathcal{T}(\Gamma))_s]) \\ &= \int d\Gamma_s \rho_\ell(\Gamma_s) e^{-\Sigma[\Gamma_s \rightarrow (\mathcal{T}(\Gamma))_s]} e^{\Sigma[\Gamma_s \rightarrow (\mathcal{T}(\Gamma))_s]} \delta(Y - \Xi[\Gamma_s \rightarrow (\mathcal{T}(\Gamma))_s]). \end{aligned} \quad (87)$$

Inserting $\mathcal{T}\mathcal{T}^{-1}$ in the above equation, we have

$$\begin{aligned} P_\tau^F(Y) &= \int d\Gamma_s \rho_\ell(\Gamma_s) \mathcal{T}\mathcal{T}^{-1} e^{-\Sigma[\Gamma_s \rightarrow (\mathcal{T}(\Gamma))_s]} e^{\Sigma[\Gamma_s \rightarrow (\mathcal{T}(\Gamma))_s]} \delta(Y - \Xi[\Gamma_s \rightarrow (\mathcal{T}(\Gamma))_s]) \\ &= \int d\Gamma_s \rho_\ell((\mathcal{T}^{-1}(\Gamma))_s) e^{-\Sigma[(\mathcal{T}^{-1}(\Gamma))_s \rightarrow \Gamma_s]} e^{\Sigma[(\mathcal{T}^{-1}(\Gamma))_s \rightarrow \Gamma_s]} \delta(Y + \Xi[(\mathcal{T}^{-1}(\Gamma))_s \rightarrow \Gamma_s]). \end{aligned} \quad (88)$$

In the last line, we used the result of the Appendix B (i.e., anti-linearity). Taking the limit and integration of $\lim_{\tau \rightarrow \infty} \int dY e^{-Y}$, we obtain

$$\begin{aligned} \lim_{\tau \rightarrow \infty} \int dY e^{-Y} P_\tau^F(Y) &= \lim_{\tau \rightarrow \infty} \int dY e^{-Y} \int d\Gamma_s \rho_\ell((\mathcal{T}^{-1}(\Gamma))_s) e^{\Sigma[(\mathcal{T}^{-1}(\Gamma))_s \rightarrow \Gamma_s]} e^{-\Sigma[(\mathcal{T}^{-1}(\Gamma))_s \rightarrow \Gamma_s]} \delta(Y + \Xi[(\mathcal{T}^{-1}(\Gamma))_s \rightarrow \Gamma_s]) \\ &= \lim_{\tau \rightarrow \infty} \int d\Gamma_s \rho_\ell((\mathcal{T}^{-1}(\Gamma))_s) e^{\Sigma[(\mathcal{T}^{-1}(\Gamma))_s \rightarrow \Gamma_s]} e^{-\Sigma[(\mathcal{T}^{-1}(\Gamma))_s \rightarrow \Gamma_s]} e^{\Xi[(\mathcal{T}^{-1}(\Gamma))_s \rightarrow \Gamma_s]} \\ &= \lim_{\tau \rightarrow \infty} Z \int d\Gamma_s \rho_{ss}(\Gamma_s) e^{-\Sigma[(\mathcal{T}^{-1}(\Gamma))_s \rightarrow \Gamma_s]} e^{\Xi[(\mathcal{T}^{-1}(\Gamma))_s \rightarrow \Gamma_s]}. \end{aligned} \quad (89)$$

Thus, we finally have

$$\begin{aligned} \lim_{\tau \rightarrow \infty} \int dY e^{-Y} P_\tau^F(Y) &= \lim_{\tau \rightarrow \infty} Z \int d\Gamma_s \rho_{ss}(\Gamma_s) e^{-\Sigma[(\mathcal{T}^{-1}(\Gamma))_s \rightarrow \Gamma_s]} e^{\Xi[(\mathcal{T}^{-1}(\Gamma))_s \rightarrow \Gamma_s]} \\ &= \lim_{\tau \rightarrow \infty} Z \left\langle \exp \left[- \sum_m \int d\mathbf{x} \int_{-\tau}^0 dt \mathbf{j}_m(\Gamma_s, \mathbf{x}; t) \cdot (\mathbf{X}_m(\mathbf{x}) - \lambda_m(\mathbf{x})) \right] \right\rangle_{ss} \\ &= \lim_{\tau \rightarrow \infty} Z \left\langle \exp \left[- \sum_m \int d\mathbf{x} \int_0^\tau dt \mathbf{j}_m(\Gamma_s, \mathbf{x}; t) \cdot (\mathbf{X}_m(\mathbf{x}) - \lambda_m(\mathbf{x})) \right] \right\rangle_{ss}. \end{aligned} \quad (90)$$

In the second line, we have used eq. (69). In the last line, we have used the stationarity of the process.

On the other hand, if we use the nonequilibrium detailed balance relation, eq. (59) to eq.(87), we have

$$\begin{aligned}
P_\tau^F(Y) &= \int d\gamma \int d\Gamma_s P_\gamma(\Gamma_s) \delta((\mathcal{T}(\Gamma))_s - \gamma) \delta(Y - \Xi[\Gamma_s \rightarrow (\mathcal{T}(\Gamma))_s]) \\
&= \int d\gamma \int d\Gamma_s P_\gamma(\Gamma_s) \delta((\mathcal{T}(\mathcal{T}(\Gamma))^*)_s - \gamma^*) e^{\Sigma[\Gamma_s \rightarrow (\mathcal{T}(\Gamma))_s]} \delta(Y - \Xi[\Gamma_s \rightarrow (\mathcal{T}(\Gamma))_s]) \\
&= \int d\gamma \int d\Gamma_s \rho_\ell(\Gamma_s) \delta((\mathcal{T}(\mathcal{T}(\Gamma))^*)_s - \gamma^*) e^{\Sigma[\Gamma_s \rightarrow (\mathcal{T}(\Gamma))_s]} \delta(Y - \Xi[\Gamma_s \rightarrow (\mathcal{T}(\Gamma))_s]) \\
&= \int d\gamma \int d\Gamma_s \rho_\ell(\Gamma_s) \delta(\Gamma_s^* - \gamma^*) e^{\Sigma[\Gamma_s \rightarrow (\mathcal{T}(\Gamma))_s]} \delta(Y - \Xi[\Gamma_s \rightarrow (\mathcal{T}(\Gamma))_s]) \\
&= \int d\gamma \rho_\ell(\gamma) e^{\Sigma[\gamma \rightarrow (\mathcal{T}(\Gamma))_s]} \delta(Y - \Xi[\gamma \rightarrow (\mathcal{T}(\Gamma))_s]).
\end{aligned} \tag{91}$$

Here we used the relation $\mathcal{T}(\mathcal{T}(\Gamma)^*) = \Gamma^*$. Then, taking the limit and integration of $\lim_{\tau \rightarrow \infty} \int dY e^{-Y}$, we have

$$\begin{aligned}
\lim_{\tau \rightarrow \infty} \int dY e^{-Y} P_\tau^F(Y) &= \lim_{\tau \rightarrow \infty} \int dY e^{-Y} \int d\gamma \rho_\ell(\gamma) e^{\Sigma[\gamma \rightarrow (\mathcal{T}(\Gamma))_s]} \delta(Y - \Xi[\gamma \rightarrow (\mathcal{T}(\Gamma))_s]) \\
&= \lim_{\tau \rightarrow \infty} \int d\gamma \rho_\ell(\gamma) e^{\Sigma[\gamma \rightarrow (\mathcal{T}(\Gamma))_s]} e^{-\Xi[\gamma \rightarrow (\mathcal{T}(\Gamma))_s]}.
\end{aligned} \tag{92}$$

Here we note that $\gamma = (\mathcal{T}(\Gamma))_s$. We obtain

$$\begin{aligned}
\lim_{\tau \rightarrow \infty} \int dY e^{-Y} P_\tau^F(Y) &= \lim_{\tau \rightarrow \infty} Z \int d\gamma \rho_{ss}(\gamma) e^{-\Xi[\gamma \rightarrow \gamma]} \\
&= \lim_{\tau \rightarrow \infty} Z \left\langle \exp \left[- \sum_m \int d\mathbf{x} \int_{-\tau}^0 dt \mathbf{j}_m(\Gamma_s, \mathbf{x}; t) \cdot \lambda_m(\mathbf{x}) \right] \right\rangle_{ss} \\
&= \lim_{\tau \rightarrow \infty} Z \left\langle \exp \left[- \sum_m \int d\mathbf{x} \int_0^\tau dt \mathbf{j}_m(\Gamma_s, \mathbf{x}; t) \cdot \lambda_m(\mathbf{x}) \right] \right\rangle_{ss}.
\end{aligned} \tag{93}$$

In the first line, we have used eq. (69). In the last line, we have used the stationarity of the process. Thus, by eqs. (90) and (93), we obtain

$$\begin{aligned}
&\lim_{\tau \rightarrow \infty} \left\langle \exp \left[- \sum_m \int d\mathbf{x} \int_0^\tau dt' \mathbf{j}_m(\Gamma_s, \mathbf{x}; t') \cdot \lambda_m(\mathbf{x}) \right] \right\rangle_{ss} \\
&= \lim_{\tau \rightarrow \infty} \left\langle \exp \left[- \sum_m \int d\mathbf{x} \int_0^\tau dt' \mathbf{j}_m(\Gamma_s, \mathbf{x}; t') \cdot (\mathbf{X}_m(\mathbf{x}) - \lambda_m(\mathbf{x})) \right] \right\rangle_{ss}.
\end{aligned} \tag{94}$$

By the definition of the cumulant generating functional, eq. (82), this gives the fluctuation theorem (i.e., Andrieux-Gaspard symmetry) $Q(\lambda; \mathbf{X}) = Q(\mathbf{X} - \lambda; \mathbf{X})$.

6. On the nonequilibrium detailed balance relation

Some readers may feel that the assumption of the nonequilibrium detailed balance relation is a key relation, and that it is "doubtful" whether it is formed. In this section, we examine the nonequilibrium detailed balance relation, eq.(59).

Back to the derivation of the fluctuation theorem for the entropy production, remind that the nonequilibrium detailed balance relation is used at eq.(65) (the McLennan-Zubarev steady state distribution), at eq.(78) (the fluctuation theorem for the entropy production), and at eq.(91) (the fluctuation theorem for the current). Suppose that the nonequilibrium detailed balance relation is given by

$$P[\gamma \rightarrow \gamma] = e^{\Theta[\gamma \rightarrow \gamma]} P[\gamma^* \rightarrow \gamma^*], \tag{95}$$

where the function $\Theta[\nu \rightarrow \gamma]$ is assumed not to be $\Sigma[\nu \rightarrow \gamma]$. Then, following the derivation of the McLennan-Zubarev steady state distribution, we obtain the following steady state distribution.

$$\rho_{ss}(\gamma) = Z_*^{-1} \rho_\ell(\gamma) \exp(\Theta[\nu \rightarrow \gamma]), \quad (96)$$

where

$$Z_* = \int d\gamma \rho_\ell(\gamma) \exp(\Theta[\nu \rightarrow \gamma]). \quad (97)$$

Further following the derivation of the fluctuation theorem for the entropy production, we have

$$\lim_{\tau \rightarrow \infty} Z_* \langle e^{-\lambda \Sigma} \rangle_{ss} = \lim_{\tau \rightarrow \infty} Z_* \langle e^{\lambda \Sigma - \Theta} \rangle_{ss}, \quad (98)$$

where

$$\langle \cdots \rangle_{ss} = \int d\gamma \rho_{ss}(\gamma) \cdots. \quad (99)$$

We find that the fluctuation theorem is not valid for the general case of $\Theta([\nu \rightarrow \gamma])$. It is clear that the fluctuation theorem is valid when the condition

$$\Theta[\nu \rightarrow \gamma] = \Sigma[\nu \rightarrow \gamma], \quad (100)$$

is formed. Thus, assuming the nonequilibrium detailed balance relation, implies the validity of the fluctuation theorem for the entropy production. The same argument for the fluctuation theorem for the current is true. Following the derivation of the fluctuation theorem for the current, we have

$$\lim_{\tau \rightarrow \infty} Z_* \langle e^{-\Theta + \Xi} \rangle_{ss} = \lim_{\tau \rightarrow \infty} Z_* \langle e^{-\Xi} \rangle_{ss}. \quad (101)$$

The fluctuation theorem for the current is valid if and only if the condition $\Theta[\nu \rightarrow \gamma] = \Sigma[\nu \rightarrow \gamma]$ consists. Thus, assuming the nonequilibrium detailed balance relation as eq.(59) implies the validity of two fluctuation theorems.

We may want to justify the nonequilibrium detailed balance relation. However, unfortunately, there is no plausible justification. We believe the analogy to similar relations derived for other types of dynamics, for instance, stochastic Markovian dynamics[28, 7]. Thus, we assume the nonequilibrium detailed balance relation, eq.(59). As a result, two fluctuation theorems hold good.

7. Nonlinear response

The cumulant generating functional for the current, i.e. eq. (82), is useful for evaluating the mean current. The current can be written in terms of thermodynamic forces, i.e., a power series of thermodynamic forces. The thermodynamic force measures the strength of the nonequilibrium property, i.e., how away from equilibrium. In order to clarify the strength of the nonequilibrium property, here we set $\mathbf{X}(\mathbf{x})$ as $\epsilon \mathbf{X}(\mathbf{x})$. The McLennan-Zubarev steady state distribution depends on the thermodynamic forces. We can expand the McLennan-Zubarev steady state distribution in a power series of thermodynamic forces. Then, the mean current is a sum of the equilibrium current ($O(\epsilon^0)$), the linear response ($O(\epsilon^1)$) (the Green-Kubo formula), and the non-linear responses ($O(\epsilon^n)$, $n \geq 2$). This idea is originally due to Gallavotti[4]. Gallavotti's idea was applied to the master equation by Andrieux and Gaspard[11, 12] to give the non-linear response. In this section, we evaluate the linear response and non-linear response in the current for the case of the McLennan-Zubarev steady distribution. We set $\mathbf{X} = \mathbf{X}'/k_B$ and $\lambda = \lambda'/k_B$. The cumulant generating functional is $Q(\lambda; \mathbf{X}) = \tilde{Q}(\lambda'; \mathbf{X}')$.

The α th component $J_\alpha(\mathbf{x})$ of the mean current at \mathbf{x} is given by the functional derivative with respect to $\lambda_\alpha(\mathbf{x})$

$$J_\alpha(\mathbf{x}) = \left. \frac{\delta Q}{\delta \lambda_\alpha(\mathbf{x})} \right|_{\lambda=0} = k_B \left. \frac{\delta \tilde{Q}}{\delta \lambda'_\alpha(\mathbf{x})} \right|_{\lambda'=0}. \quad (102)$$

By definition, we have

$$J_\alpha(\mathbf{x}) = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t dt' \langle j_\alpha(\Gamma, \mathbf{x}; t') \rangle_{ss}. \quad (103)$$

The mean current is expanded in thermodynamic forces.

$$\begin{aligned}
J_\alpha(\mathbf{x}) &= J_\alpha^{(\text{eq})}(\mathbf{x}) + \frac{\epsilon}{1!} \sum_\beta \int d\mathbf{x}_1 \mathcal{L}_{\alpha\beta}(\mathbf{x}, \mathbf{x}_1) X'_\beta(\mathbf{x}_1) \\
&+ \frac{\epsilon^2}{2!} \sum_{\beta, \gamma} \iint d\mathbf{x}_1 d\mathbf{x}_2 \mathcal{M}_{\alpha\beta\gamma}(\mathbf{x}, \mathbf{x}_1, \mathbf{x}_2) X'_\beta(\mathbf{x}_1) X'_\gamma(\mathbf{x}_2) \\
&+ \frac{\epsilon^3}{3!} \sum_{\beta, \gamma, \delta} \iiint d\mathbf{x}_1 d\mathbf{x}_2 d\mathbf{x}_3 \mathcal{N}_{\alpha\beta\gamma\delta}(\mathbf{x}, \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3) X'_\beta(\mathbf{x}_1) X'_\gamma(\mathbf{x}_2) X'_\delta(\mathbf{x}_3) \\
&+ \dots,
\end{aligned} \tag{104}$$

where

$$\begin{aligned}
J_\alpha^{(\text{eq})}(\mathbf{x}) &= \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t dt' \langle j_\alpha(\mathbf{x}; t') \rangle_{\text{eq}} \\
&= \int d\Gamma \rho_{\text{can}}(\Gamma) j_\alpha(\Gamma, \mathbf{x}; 0) \\
&= \langle j_\alpha(\Gamma, \mathbf{x}; 0) \rangle_{\text{eq}}.
\end{aligned} \tag{105}$$

$\langle \dots \rangle_{\text{eq}}$ is the average over the canonical ensemble. Here we consider that the system in local equilibrium is near equilibrium, that is, the temperature field is almost constant and $\mathbf{u} \approx 0$. So ρ_ℓ is replaced by ρ_{can} . We have used the fact that the time-dependence is washed out in the average over the canonical ensemble, i.e., the average over the equilibrium state.

On the other hand, we obtain the mean current by averaging over the McLennan-Zubarev steady state distribution.

$$J_\alpha(\mathbf{x}) = \langle j_\alpha(\Gamma, \mathbf{x}; 0) \rangle_{\text{ss}} = \int d\Gamma \rho_{\text{ss}}(\Gamma) j_\alpha(\Gamma, \mathbf{x}; 0). \tag{106}$$

Expanding $\rho_{\text{ss}}(\Gamma)$ in terms of thermodynamic forces, we also obtain the mean current in a power series of thermodynamic forces. Of course, each term in two expansions from eqs. (102) and (106) should coincide identically.

7.1. Linear response

The function $\mathcal{L}_{\alpha\beta}(\mathbf{x}, \mathbf{x}_1)$ is given by

$$\mathcal{L}_{\alpha\beta}(\mathbf{x}, \mathbf{x}_1) = k_B \frac{\delta^2 \tilde{Q}}{\delta \lambda'_\alpha(\mathbf{x}) \delta X'_\beta(\mathbf{x}_1)}(\mathbf{0}; \mathbf{0}) = \frac{1}{k_B} \frac{\delta^2 Q}{\delta \lambda_\alpha(\mathbf{x}) \delta X_\beta(\mathbf{x}_1)}(\mathbf{0}; \mathbf{0}). \tag{107}$$

Differentiating eq. (84), we obtain

$$\begin{aligned}
\frac{\delta^2 Q}{\delta \lambda_\alpha(\mathbf{x}) \delta X_\beta(\mathbf{x}_1)}(\lambda; \mathbf{X}) &= -\frac{\delta^2 Q}{\delta \lambda_\alpha(\mathbf{x}) \delta \lambda_\beta(\mathbf{x}_1)}(\mathbf{X} - \lambda; \mathbf{X}) \\
&- \frac{\delta^2 Q}{\delta \lambda_\alpha(\mathbf{x}) \delta X_\beta(\mathbf{x}_1)}(\mathbf{X} - \lambda; \mathbf{X}).
\end{aligned} \tag{108}$$

Therefore, setting $\lambda = \mathbf{0}$ and $\mathbf{X} = \mathbf{0}$, we have

$$\mathcal{L}_{\alpha\beta}(\mathbf{x}, \mathbf{x}_1) = \frac{1}{k_B} \frac{\delta^2 Q}{\delta \lambda_\alpha(\mathbf{x}) \delta X_\beta(\mathbf{x}_1)}(\mathbf{0}; \mathbf{0}) = -\frac{1}{2k_B} \frac{\delta^2 Q}{\delta \lambda_\alpha(\mathbf{x}) \delta \lambda_\beta(\mathbf{x}_1)}(\mathbf{0}; \mathbf{0}). \tag{109}$$

This gives the Onsager reciprocity relation[31, 32]

$$\mathcal{L}_{\alpha\beta}(\mathbf{x}, \mathbf{x}_1) = \mathcal{L}_{\beta\alpha}(\mathbf{x}_1, \mathbf{x}). \tag{110}$$

Expanding the McLennan-Zubarev steady state distribution in thermodynamic forces, namely expanding the exponential function and the normalization constant Z , we get

$$\mathcal{L}_{\alpha\beta}(\mathbf{x}, \mathbf{x}_1) = \frac{1}{2k_B} \int_{-\infty}^{\infty} dt \phi_{\alpha\beta}^{(2)}(\mathbf{x}, \mathbf{x}_1; t), \quad (111)$$

where $\phi_{\alpha\beta}^{(2)}(\mathbf{x}, \mathbf{x}_1; t)$ is the two-point current correlation function,

$$\phi_{\alpha\beta}^{(2)}(\mathbf{x}, \mathbf{x}_1; t) = \langle [j_\alpha(\Gamma, \mathbf{x}; t) - \langle j_\alpha(\Gamma, \mathbf{x}; t) \rangle_{\text{eq}}] [j_\beta(\Gamma, \mathbf{x}_1; 0) - \langle j_\beta(\Gamma, \mathbf{x}_1; 0) \rangle_{\text{eq}}] \rangle_{\text{eq}}. \quad (112)$$

Here note that the average is taken over the local equilibrium state.

Comparing the term in eq. (102) with that of eq. (106) in the order of $O(\epsilon^1)$,

$$\int_0^\infty dt \phi_{\alpha\beta}^{(2)}(\mathbf{x}, \mathbf{x}_1; t) = \int_0^\infty dt \phi_{\alpha\beta}^{(2)}(\mathbf{x}, \mathbf{x}_1; -t). \quad (113)$$

By eq. (113), $\mathcal{L}_{\alpha\beta}(\mathbf{x}, \mathbf{x}_1)$ is given by

$$\mathcal{L}_{\alpha\beta}(\mathbf{x}, \mathbf{x}_1) = \frac{1}{k_B} \int_0^\infty dt \phi_{\alpha\beta}^{(2)}(\mathbf{x}, \mathbf{x}_1; t). \quad (114)$$

Equation (114) is nothing but the Green-Kubo formula [33, 34, 35].

7.2. Non-linear response of $O(\epsilon^2)$

The function $\mathcal{M}_{\alpha\beta\gamma}(\mathbf{x}, \mathbf{x}_1, \mathbf{x}_2)$ is given by

$$\mathcal{M}_{\alpha\beta\gamma}(\mathbf{x}, \mathbf{x}_1, \mathbf{x}_2) = \frac{1}{k_B^2} \frac{\delta^3 \mathcal{Q}}{\delta \lambda_\alpha(\mathbf{x}) \delta X_\beta(\mathbf{x}_1) \delta X_\gamma(\mathbf{x}_2)}(\mathbf{0}; \mathbf{0}). \quad (115)$$

In the same way of the previous subsection, differentiating eq. (84), and setting λ and $\mathbf{X} = 0$, we obtain

$$\begin{aligned} \mathcal{M}_{\alpha\beta\gamma}(\mathbf{x}, \mathbf{x}_1, \mathbf{x}_2) &= -\frac{1}{2k_B^2} \frac{\delta^3 \mathcal{Q}}{\delta \lambda_\alpha(\mathbf{x}) \delta \lambda_\beta(\mathbf{x}_1) \delta \lambda_\gamma(\mathbf{x}_2)}(\mathbf{0}; \mathbf{0}) \\ &\quad -\frac{1}{2k_B^2} \frac{\delta^3 \mathcal{Q}}{\delta \lambda_\alpha(\mathbf{x}) \delta \lambda_\beta(\mathbf{x}_1) \delta X_\gamma(\mathbf{x}_2)}(\mathbf{0}; \mathbf{0}) \\ &\quad -\frac{1}{2k_B^2} \frac{\delta^3 \mathcal{Q}}{\delta \lambda_\alpha(\mathbf{x}) \delta \lambda_\gamma(\mathbf{x}_1) \delta X_\beta(\mathbf{x}_2)}(\mathbf{0}; \mathbf{0}). \end{aligned} \quad (116)$$

Evaluating the right hand side of eq. (116), it leads

$$\begin{aligned} &\mathcal{M}_{\alpha\beta\gamma}(\mathbf{x}, \mathbf{x}_1, \mathbf{x}_2) \\ &= \frac{1}{2k_B^2} \left\{ \int_{-\infty}^{\infty} dt_1 \int_{-\infty}^0 dt_2 + \int_{-\infty}^0 dt_1 \int_{-\infty}^{\infty} dt_2 \right\} \phi_{\alpha\beta\gamma}^{(3)}(\mathbf{x}, \mathbf{x}_1, \mathbf{x}_2; t_1, t_2), \end{aligned} \quad (117)$$

where

$$\begin{aligned} \phi_{\alpha\beta\gamma}^{(3)}(\mathbf{x}, \mathbf{x}_1, \mathbf{x}_2; t_1, t_2) &= \langle j_\alpha(\Gamma, \mathbf{x}; 0) j_\beta(\Gamma, \mathbf{x}_1; t_1) j_\gamma(\Gamma, \mathbf{x}_2; t_2) \rangle_{\text{eq}} - \langle j_\alpha(\Gamma, \mathbf{x}; 0) j_\beta(\Gamma, \mathbf{x}_1; t_1) \rangle_{\text{eq}} \langle j_\gamma(\Gamma, \mathbf{x}_2; t_2) \rangle_{\text{eq}} \\ &\quad - \langle j_\alpha(\Gamma, \mathbf{x}; 0) \rangle_{\text{eq}} \langle j_\beta(\Gamma, \mathbf{x}_1; t_1) j_\gamma(\Gamma, \mathbf{x}_2; t_2) \rangle_{\text{eq}} - \langle j_\alpha(\Gamma, \mathbf{x}; 0) j_\gamma(\Gamma, \mathbf{x}_2; t_2) \rangle_{\text{eq}} \langle j_\beta(\Gamma, \mathbf{x}_1; t_1) \rangle_{\text{eq}} \\ &\quad + 2 \langle j_\alpha(\Gamma, \mathbf{x}; 0) \rangle_{\text{eq}} \langle j_\beta(\Gamma, \mathbf{x}_1; t_1) \rangle_{\text{eq}} \langle j_\gamma(\Gamma, \mathbf{x}_2; t_2) \rangle_{\text{eq}}. \end{aligned} \quad (118)$$

On the other hand, we obtain the term in $O(\epsilon^2)$ from eq. (106). This should coincide with eq. (117). This comparison gives a non-trivial relation of the three-point current correlation function

$$\left\{ \int_0^\infty dt_1 \int_{-\infty}^0 dt_2 + \int_{-\infty}^0 dt_1 \int_0^\infty dt_2 \right\} \phi_{\alpha\beta\gamma}^{(3)}(\mathbf{x}, \mathbf{x}_1, \mathbf{x}_2; t_1, t_2) = 0. \quad (119)$$

With this relation, $\mathcal{M}_{\alpha\beta\gamma}(\mathbf{x}, \mathbf{x}_1, \mathbf{x}_2)$ is given by

$$\mathcal{M}_{\alpha\beta\gamma}(\mathbf{x}, \mathbf{x}_1, \mathbf{x}_2) = \frac{1}{k_B^2} \int_{-\infty}^0 dt_1 \int_{-\infty}^0 dt_2 \phi_{\alpha\beta\gamma}^{(3)}(\mathbf{x}, \mathbf{x}_1, \mathbf{x}_2; t_1, t_2). \quad (120)$$

The non-trivial relations of eqs. (113) and (119) are the consequence of the time-reversal symmetry, i.e., the fluctuation theorem $Q(\lambda; \mathbf{X}) = Q(\mathbf{X} - \lambda; \mathbf{X})$. In the same way as shown above, non-trivial relations would be obtained for the order of $O(\epsilon^n)$ ($n \geq 3$).

8. Concluding remarks

In this paper, we have derived the McLennan-Zubarev steady state distribution from the nonequilibrium detailed balance relation, and have proved the fluctuation theorems for the entropy production and for the currents. As shown in the derivation and the proof, the McLennan-Zubarev steady state distribution and the fluctuation theorems have the same root, in the sense that they are derived from the same relations, i.e., the time-reversal symmetry and the nonequilibrium detailed balance relation.

As a consequence of the fluctuation theorem for the currents, we have derived the expression for the mean current in terms of thermodynamic forces. The linear response in the order $O(\epsilon^1)$ and non-linear response in the order of $O(\epsilon^2)$ are derived. The linear response exactly coincides with the Green-Kubo formula. The non-linear response of $O(\epsilon^2)$ gives the first correction to the Green-Kubo formula. We have also obtained the non-trivial relations of the correlation functions for the orders of $O(\epsilon^1)$ and $O(\epsilon^2)$. These non-trivial relations of the correlation functions are a consequence of time-reversal symmetry. The corrections $O(\epsilon^n)$, $n \geq 3$ can be derived in a systematic way [12]. In addition, the results of this paper are exact without approximations except replacing the local equilibrium state by the canonical distribution for the section of the nonlinear response.

Acknowledgements

The author thanks Eiji Konishi for careful reading the manuscript. This work is supported by JSPS KAKENHI (No.24654122).

Appendix A

In this appendix, we show McLennan's derivation of the McLennan-Zubarev steady state distribution[14]. We consider the system which is coupled with baths. Now the Hamiltonian is given by

$$H_{\text{tot}} = H_S + H_B + U, \quad (A. 1)$$

where H_S is the Hamiltonian of the system, H_B is the Hamiltonian of the baths, and U is the interaction between the system and baths. U does not include the momenta. The whole system is a Hamiltonian system. The Liouville equation is

$$\frac{\partial f}{\partial t} = \{H_{\text{tot}}, f\}. \quad (A. 2)$$

We define the variables. The position and momentum coordinates for the whole system is $\Gamma = (\Gamma_s, \Gamma_b)$, where Γ_s and Γ_b is the coordinates for the systems and the baths, respectively. We set $f = \rho\chi\eta$. ρ and $\chi\eta$ are

$$\rho = \int d\Gamma_b f, \quad \chi\eta = \int d\Gamma_s f. \quad (A. 3)$$

The normalization conditions are

$$\int d\Gamma_s \rho = 1, \quad \int d\Gamma_b \chi\eta = 1. \quad (A. 4)$$

Integrating eq. (A. 2) with respect to the bath coordinates, we obtain

$$\frac{\partial \rho}{\partial t} + \{\rho, H_s\} + \int d\Gamma_b \{f, H_b\} + \int d\Gamma_b \{f, U\} = 0. \quad (\text{A. 5})$$

The third term is the divergence. Thus, this is zero if the contribution vanishes sufficiently rapidly toward the surface of the system. This point is very similar to the case in Zubarev's derivation. Here we omit the term which appears in the use of the Gauss theorem. Then, we get

$$\frac{\partial \rho}{\partial t} + \{\rho, H_s\} + \frac{\partial(\rho F_\alpha)}{\partial p_\alpha} = 0. \quad (\text{A. 6})$$

Here the force F_α is

$$F_\alpha = - \int d\Gamma_b \chi \eta \frac{\partial U}{\partial q_\alpha}. \quad (\text{A. 7})$$

This force is not conservative. If one set

$$\frac{\partial F_\alpha}{\partial p_\alpha} = - \int d\mathbf{S} \cdot \mathbf{J}_S(\mathbf{x}; t), \quad (\text{A. 8})$$

then one obtains the McLennan-Zubarev steady state distribution ρ_{ss} . Here $\mathbf{J}_S(\mathbf{x}; t)$ is

$$\mathbf{J}_S(\mathbf{x}; t) = \beta(\mathbf{x}; t)(\mathbf{J}_H(\Gamma, \mathbf{x}) - \mu^* \mathbf{j}(\Gamma, \mathbf{x}) - \mathbf{u}(\mathbf{x}) \cdot \mathbf{T}(\Gamma, \mathbf{x})). \quad (\text{A. 9})$$

This is the way of McLennan's derivation.

Appendix B

In this appendix, we note a character of the Liouvillian. Remember characters of the Liouville equation.

$$\frac{\partial \rho}{\partial t} = \mathcal{L}\rho, \quad (\text{B. 1})$$

where

$$\mathcal{L} = \{H, \} = \sum_{i=1}^N \left\{ \frac{\partial H}{\partial \mathbf{q}_i} \cdot \frac{\partial}{\partial \mathbf{p}_i} - \frac{\partial H}{\partial \mathbf{p}_i} \cdot \frac{\partial}{\partial \mathbf{q}_i} \right\}. \quad (\text{B. 2})$$

is the Liouvillian. The formal solution of the distribution is

$$\rho(\Gamma, t) = e^{\mathcal{L}t} \rho(\Gamma, 0). \quad (\text{B. 3})$$

The coordinates is evolved as

$$\Gamma(t) = e^{-\mathcal{L}t} \Gamma(0). \quad (\text{B. 4})$$

The Liouvillian has an anti-linearity.

$$\int d\Gamma F(\Gamma) [\mathcal{L}G(\Gamma)] = - \int d\Gamma [\mathcal{L}F(\Gamma)] G(\Gamma). \quad (\text{B. 5})$$

This anti-linearity implies the following.

$$\begin{aligned} \int d\Gamma F(\Gamma) G(\mathcal{T}(\Gamma)) &= \int d\Gamma F(\Gamma) \mathcal{T} G(\Gamma) \\ &= \int d\Gamma G(\Gamma) \mathcal{T}^{-1} F(\Gamma) \\ &= \int d\Gamma G(\Gamma) F(\mathcal{T}^{-1}(\Gamma)). \end{aligned} \quad (\text{B. 6})$$

Here we set $\mathcal{T} = e^{-\mathcal{L}\tau}$ which is a time-evolution operator for the coordinates. This identity is used in the text.

References

- [1] Evans, D. J., Cohen, E. G. D., and Morriss, G. P.: Probability of Second Law Violations in Shearing Steady States. *Phys. Rev. Lett.* **71** 2401 (1993).
- [2] Gallavotti, G. and Cohen, E. G. D.: Dynamical Ensembles in Nonequilibrium Statistical Mechanics. *Phys. Rev. Lett.* **74** 2694 (1995).
- [3] Gallavotti, G. and Cohen, E. G. D.: Dynamical Ensembles in Stationary States. *J. Stat. Phys.* **80** 931 (1995).
- [4] Gallavotti, G.: Extension of Onsager's Reciprocity to Large Fields and the Chaotic Hypothesis. *Phys. Rev. Lett.* **77** 4334 (1996).
- [5] Kurchan, J.: Fluctuation Theorem for Stochastic Dynamics. *J. Phys. A* **31** 3719 (1998).
- [6] Lebowitz, J. L. and Spohn, H.: A Gallavotti-Cohen-type Symmetry in the Large Deviation Functional for Stochastic Dynamics. *J. Stat. Phys.* **95** 333 (1999).
- [7] Crooks, G. E.: Entropy Production Fluctuation Theorem and the Nonequilibrium Work Relation for Free Energy Differences. *Phys. Rev. E* **60** 2721 (1999).
- [8] Jarzynski, C.: Hamiltonian Derivation of a Detailed Fluctuation Theorem. *J. Stat. Phys.* **98** 77 (2000).
- [9] D. J. Evans and D. J. Searles: The fluctuation theorem *Adv. Phys.* **51** 1529 (2002).
- [10] Gaspard, P.: Fluctuation Theorem for Nonequilibrium Reactions. *J. Chem. Phys.* **120** 8898 (2004).
- [11] Andrieux, D. and Gaspard, P.: Fluctuation Theorem and Onsager Reciprocity Relations. *J. Chem. Phys.* **121** 6167 (2004).
- [12] Andrieux, D. and Gaspard, P.: A Fluctuation Theorem for Currents and Non-linear Response Coefficients. *J. Stat. Mech.* P02006 (2007).
- [13] Andrieux, D. and Gaspard, P.: Fluctuation Theorem for Currents and Schnakenberg Network Theory. *J. Stat. Phys.* **127** 107 (2007).
- [14] McLennan, J. A.: Statistical Mechanics of the Steady State. *Phys. Rev.* **115** 1405 (1959).
- [15] McLennan, J. A.: The Formal Statistical Theory of Transport Processes. *Adv. Chem. Phys.* **5** 261 (1963).
- [16] McLennan, J. A.: Introduction to Nonequilibrium Statistical Mechanics. (Prentice Hall, Englewood Cliffs, 1990).
- [17] Zubarev, D. N.: Statistical Operator for Non-equilibrium Systems. *Doklady Akademii Nauk SSSR* **143** 74 (1962).
- [18] Zubarev, D. N.: Nonequilibrium Statistical Thermodynamics, (Consultants Bureau, New York, 1974).
- [19] Sano, M. M.: The Steady State Distribution of the Master Equation. *J. Phys. A* **41** 435001 (2008).
- [20] Maes, C. and Netočný, K.: Rigorous Meaning of McLennan Ensembles. *J. Math. Phys.* **51** 015219 (2010).
- [21] Colangeli, M., Maes, C., and Wynants, B.: A Meaningful Expansion Around Detailed Balance. *J. Phys. A* **44** 095001 (2011).
- [22] Komatsu, T. S. and Nakagawa, N.: Expression for the Stationary Distribution in Nonequilibrium Steady States. *Phys. Rev. Lett.* **100** 030601 (2008).
- [23] Komatsu, T. S., Nakagawa, N., Sasa, S. -I., and Tasaki, H.: Steady-state Thermodynamics for Heat Conduction: Microscopic derivation. *Phys. Rev. Lett.* **100** 230602 (2008).
- [24] Komatsu, T. S., Nakagawa, N., Sasa, S. -I., and Tasaki, H.: Representation of Nonequilibrium Steady States in Large Mechanical Systems. *J. Stat. Phys.* **134** 401 (2009).
- [25] Komatsu, T. S., Nakagawa, N., Sasa, S. -I. and Tasaki, H.: Entropy and Nonlinear Nonequilibrium Thermodynamic Relation for Heat Conducting Steady States. *J. Stat. Phys.* **142** 127 (2011).
- [26] Tasaki, S. and Matsui, T.: Fluctuation theorem, Nonequilibrium Steady States and MacLennan-Zubarev Ensembles of a Class of Large Quantum Systems. in *Fundamental Aspects of Quantum Physics*, Edited by L. Accardi and S. Tasaki (World Scientific, Singapore, 2003), p.100.
- [27] Tasaki, S. and Matsui, T.: Note on MacLennan-Zubarev Ensembles and QuasiStatic Processes. *RIMS Kôkyûroku* **1507** 118 (2006).
- [28] Crooks, G. E.: Nonequilibrium Measurements of Free Energy Differences for Microscopically Reversible Markovian Systems *J. Stat. Phys.* **90**, 1481 (1998).
- [29] de Groot S. R. and Mazur, P.: *Non-equilibrium Thermodynamics* (Dover, New York, 1984)
- [30] K. Kitahara and K. Yoshikawa: *Science of Nonequilibrium Systems I: Phenomenology of reactions, diffusion, and convection* (Koudansha scientific, Tokyo, 1994) (in Japanese).
- [31] Onsager, L.: Reciprocal Relations in Irreversible Processes. I. *Phys. Rev.* **37** 405 (1931)
- [32] Onsager, L.: Reciprocal Relations in Irreversible Processes. II. *Phys. Rev.* **38** 2265 (1931)
- [33] Green, M. S.: Markoff Random Processes and the Statistical Mechanics of Time-Dependent Phenomena. *J. Chem. Phys.* **20** 1281 (1952)
- [34] Green, M. S.: Markoff Random Processes and the Statistical Mechanics of Time-Dependent Phenomena II Irreversible Processes in Fluids *J. Chem. Phys.* **22** 398 (1954)
- [35] Kubo, R.: Statistical Mechanical Theory of Irreversible Processes I.: General Theory and Simple Applications to Magnetic and Conduction Problems. *J. Phys. Soc. Jpn* **12** 570 (1957).